DOCUMENT RESUME

ED 209 354 TM 810 696

AUTHOR Russell, Elbert W.
TITLE Memory Testing: The Next Step.
PUB DATE 5 Sep 80
NOTE 11p.; Paper presented at the Annual Meeting of the American Psychological Association (Montreal, Canada, September, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Memory; Neurology; Psychological Testing; Psychology; *Psychometrics; *Test Construction; Theories
IDENTIFIERS *Neuropsychology

ABSTRACT

Neuropsychology combines psychometrics and the study of brain processes. There are two approaches to neuropsychology. The neurologic approach uses brief tests flexibly. The psychometric approach uses standardized graded scales in a set battery. Cortical testing in the form of intellectual abilities testing is 70 years old, yet psychometric principles have never been adequately applied in neuropsychology. Psychometric principles of validity, reliability, scale homogeneity, redundancy, coverage and utility need to be applied. The Wechsler Memory Scale and Luria Nebraska Battery are inadequate in most of these areas. The Revised Wechsler Memory Scale overcomes some of the problems. However, a new memory test derived from memory theory needs to be developed. Types of memory that need to be included are immediate memory, consolidation, remote, verbal, figural, episodic, learning, rote and meaningful memory. Experimental tests of each type are available. They must simply be properly selected and placed into a battery. Psychometric principles should then be applied. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED209254

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

MEMORY TESTING: THE NEXT STEP

Elbert W. Russell

Veterans Administration Medical Center
Miami, Florida

Paper presented at the 88th Annual Convention of the American Psychological Association 1980, Montreal, Canada.

Neuropsychologists are familiar with the definition of neuropsychology as the study of brain behavior relationships. This definition has been utilized by most books on neuropsychology in one form or another. However, from the point of view of psychology the definition of neuropsychology might be more specifically stated as the application of psychometric methods to the study of brain processes. This may appear to be another radical and in some quarters, unpopular, definition of neuropsychology. However, it does possess the advantage that it states the unique relationship between psychology and brain processes. The particular contribution of psychology to the neurosciences is the mental test. Mental or psychological test development has been one of the major areas of clinical psychlology since its inception near the turn of the century. Neurology, biology and some other sciences also study brain-behavior relationships and each has its own particular set of instruments and methods. It is psychology that makes maximum use of and understands the use and construction of mental or psychological tests.

In one sense the entire area of neuropsychology is based on a relatively obvious but seldom stated principle that "a mental function can not be studied until there is a test for it." That is, until there is some way of isolating and measuring, even informally, a mental function it can not be studied and often it is not even known to exist. A test may be informal such as the early tests for aphasia. The relatively early understanding of aphasia was largely due to the situation that when speech is impaired by brain damage, it is obvious and can be informally tested. Impairment of other functions is not so obvious and consequently the understanding of such functions required the development of specific tests. The understanding of right hemisphere functions, which was quite recent, depended upon the construction of new psychological mental testing methods that could be utilized to demonstrate the effects of a severed corpus callosum. Although for a considerable time neurologists were aware of people with a congenital lack or agenisis of corpus callosum, no one had been able to demonstrate that the lack of corpus callosum affected their behavior until the proper psychological testing methods were constructed.

Mental tests obviously are the foundation for both experimental and clinical studies. Almost every experimental study requires tests in order to demonstrate that some change has been produced by the experimental conditions. Part of genius of such experimentalists as Milner and Butters is due to their ingenuity in developing new tests. It is also obvious that clinical assessment in neuropsychology utilizes tests. In fact there is little that the neuropsychologist can contribute to the assessment of brain damage without clincial tests. The WAIS which is probably the best

TM 810 652

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
E. W. Russell
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

constructed general mental test used in neuropsychology, is the only test universally used in neuropsychology. Even neuropsychologists such as Luria (1973,1980) who disparaged psychometics utilizes their own version of tests. Every demonstration of a brain function that Luria has made was based on one or another test.

There are at present two approaches to neuropsychology testing. One is derived from neurology. This neurological approach utilizes many brief plus-minus tests. The primary champion of this approach is Luria (Luria & Majovski, 1977). In fact Luria was primarily a neurologist who utilized neurological methods to study higher cortical functions rather than subcortical processes.

The second approach to neuropsychology is the psychometric approach which utilizes standardized graded scales. Most of the type of the neuropsychology testing which is done in the United States in one way or another use psychometric tests that have graded scales. Psychologist in the United States have in fact been testing cortical functions for almost seventy years. The entire intelligence test and abilities testing area of psychology began just after the turn of the century with the development of the Binet scales. Consequently psychologists are not novices at testing cortical functions even though they have only fairly recently applied their cortical testing abilities to brain pathology.

The area of psychometrics has become highly developed both in development of new tests and the principals utilized for designing and constructing tests. Unfortunately neuropsychology has not utilized this development. The level of sophistication in neuropsychology is not much greater than that which was prevealent in the 1940's for psychometrics. Such psychologists as Guilford (1976), Cattel (1943) and Horn (1970) have developed psychometric methods and testing to a high degree during the thirty years or more since World War II. This information and sophistication has not been applied to neurospsychological testing. As a matter of fact a recent book by Kline (1979) which summarizes the work done by psychometrics in this country does not mention neuropsychological testing. This in itself is an inditement of neuropsychologists' use of measurement methods. Secondly, in terms of test construction principles, only two tests in the area of neuropsychology have been adequately developed. One of these is the Wechsler Adult Intelligence Scale which was, of course, not developed as a neuropsychological test but which has become the only universally used neuropsychological test. The second test is the Boston Aphasia Examination (Goodglass & Kaplan, 1976) which was developed specifically for the limited area of aphasia. The Boston Aphasia Test appears to be the only adequately constructed test specifically designed for the area of neuropsychology. The other test batteries which are popular such as the Halstead-Reitan Battery (HRB) (Boll, 1981) and the Luria-Nebraska Battery (LNB) (Golden, 1981) have in one way or another had glaring deficiences. The HRB has never been adequately standardized. The LNB has scales which are not homogenious (Nunnally, 1967) that is the scales of the LNB are measuring a great many different items within each scale. Thus, the scales are largely meaningless. Also the LNB does not use graded scales for each item tested.

Principles of Test Construction

There are a number of general principles of test construction that apply to neuropsychological tests as well as to other tests in psychology. The concepts of reliability and validity are so well known that little needs to be said about them in this relatively short presentation. It should be noted that very few of the neuropsychological tests that are presently being utilized have been tested for reliability. Reliability can not be the test-retest type of course since neuropsychological conditions are often unstable. Consequently the type of reliability will have to be internal rather than test-retest reliability.

In regard to validity most neuropsychological tests have a criterion for validity which is some form of brain damage, either location of damage or of a particular type of diagnosis. While this will remain the major form of validation other forms of validity such as construct validity might be utilized to a greater extent than they have been. Factor analysis is helpful in construct validity and it has been utilized in the construction of the Boston Aphasia Examination. This appears to be the only instance that construct validity has been utilized in the development of neuropsychological tests. The HRB has been factored analyzed a number of times and its factor structure is fairly well known. However the analyses have not been utilized to construct either a modification of the HRB or the consruction of new tests.

One of the primary requirements for the construction of neuropsychological tests or any tests is homogeneity of items (Nunnally, 1967). Guilford (1954) calls this "univocality." That is the items in a scale should all be measuring the same thing if the scale is to mean anything (Lezak, 1976, p 119). Homogeneity was so obvious that the question as to homogeneity of a scale is seldom raised. This has only become a question since it is evident that the scales in the LNB are not homogenious. The primary method of determining homogeneity of a scale is through a factor analysis (Nunnally, 1967). A homogenious scale will have only one major factor. According to a study done by Hammele (1979) all the scales in the LNB measure more than one factor. These were as follows: Motor - seven factors; Rhythm - two factors; Tactile - two factors; Visual - three factors; Receptive Speech - two factors; Expressive Speech - five factors, Writing - two factors; Reading - two factors; Arithmatic - two factors; Memory - two factors; Intelligence - four factors.

These scales have items measuring as many as seven different functions. All scales have at least two factors. Consequently it is not at all clear what the particular scales in the LNB are measuring.

Another principle is that of graded scales. psychometric scales in general are graded in terms of difficulty or if the items are all relatively the same difficulty then in terms of the number of items that are correct. The items in the LNB at most have three steps. Thus, the items do not form graded scales. The concept of graded scales is central to psychometric testing and consequently in most testing it is simply being taken for granted.

Another concept that is not often referred to in neuropsychological testing batteries is that of redundancy. Batteries particularly in neuropsychology should have redundancy built into them. That is, it is

always possible due to any number of antifactual reasons for a person to do poorly on a particular scale when in fact that function is not impaired. Consequently some form of cross reference or redundancy should be built in neuropsychologicl tests so that scales can be crossed checked with each other.

The coverage of brain functions by a battery is of course, extremely important. This is particularly important if a set battery is utilized. The coverage should be adequate in terms of types of mental functions and in terms of localization of functions. Since localization and functions go together when one is dealing with elementary functions these can often be handled simultaneously. The concept of coverage has only recently become a requirement for neuropsychological testing, since it is only recently that the theory has developed such that various tests can be classified in terms of functions in an adequate manner.

Usage is of course another important requirement. This is particularly true for tests that are not in set batteries. Tests should be designed to be related to a particular use. Set batteries themselves should be designed in terms of general usage so that a broad range of uses can be met by a neurospychological battery. Usage is often related to validity in that the criterion for validity may be the particular use for which a test was designed to assess. One instance of an error in this requirement was the original Weschler Memory Scale. The test was supposed to be a measure of memory however the criterion used to validate was whether it could separate brain damage from normals. Even though the test was supposedly used as a test of memory this use was not applied as a criterion.

Finally, there is the important concept of efficiency. Neuropsychological tests should be as short as possible in terms of time and cost. However, it should be noted here that a short test is not necessarily the most efficient if it will not provide you with the information that is necessary to evaluate or assess the patient.

Memory Tests - Critique

As has been pointed out (Erickson & Scott, 1977) the Wechsler Memory Scale (WMS) (Wechsler, 1945) is the only major memory test in the area of clinical neuropsychology. The reviews of this test have demonstrated that it is deficient in almost every respect (Prigatano, 1977, 1978). This can be expected since it developed 35 years ago. At that time the theory of memory was almost nonexistent and the fact that the Wechsler Memory scale has held up as well as it has is a tribute to Wechsler's intuition.

In terms of some of the principles which have just been described the WMS has neither good reliability nor validity. Prigatano (1977) has examined these relationships in the Wechsler Memory Scale and found it wanting in terms of validity. One primary problem is that the scale was not validated on measures of memory but on whether or not it was effected by brain damage. This is an important variable, of course, for any neuropsychological test but it is not the proper criterion to validate a memory test.

Secondly, the Wechsler Memory Scale as it was originally designed was not univocal. It is heterogeneous in that it combines a number of types of memory tests into one scale which is labeled "memory." The orientation section is not a direct measure of memory at all. At least four types of

memory are being measured by the rest of the memory scale. The problem with combining a number of items into a single scale of memory is that failure in one area will reduce the total score without indicating which type of memory is impaired. The Wechsler Memory Scale for instance has been shown to be largely a left hemisphere test (Prigatano, 1978). Examination demonstrates that three out of four tests are left hemisphere tests. The figural memory test is a right hemisphere test (Russell, 1975) but its effects are dominated by the other tests that were verbal and left hemisphere tests.

In order to overcome some of these problems Russell (1975) revised the Wechsler Memory Scale. He set up separate scoring categories for two of the subscales in the WMS thus making them homogenious. The Digit Span in the original WAIS is also a univocal test. In terms of coverage a factor analysis of the revised WMS utilized in conjunction with the Halstead-Reitan battery and the WAIS demonstrated that five types of memory were being measured (Russell, 1978). Thus its coverage was considerably enlarged. The validity and reliability of the revised WMS were tested and appeared adequate (Russell, 1975). Subsequently it has been revalidated (Logue & Wyrich, 1979).

The Luria-Nebraska Battery has a memory scale incorporated into the test. However, this scale is heterogeneous and while one score is obtained, inspection of the scale shows that it is made up of a considerable number of individual types of memory. This analysis found that one item is not a memory item; Four of the items measure immediate verbal memory, three measure immediate nonverbal memory, one measures the mixture of immediate and long term memory, another measures immediate and long term figural memory and finally three items measure long term learning. Low scores on this test do not indicate the types of memory that are being impaired. The test appears to be more sensitive to immediate memory than to long term memory. Several studies have found that immediate memory and long term memory are almost uncorrelated.

Memory Theory

Recent years have witnessed the great development of theory in the area of memory. At the present time almost a book a year is published on memory. As a by-product of this resurgence and interest in theory and memory a large number of experimental tests of memory have been developed. In the area of psychometrics (1967) Guilford has produced a large number of additional memory tests and demonstrated some of the relationships between these memory tests and other cognitive functions.

Neither the tests which have been created by experimental work in memory or the memory measures developed by psychometrics have been applied to the clinical measurement of memory impaired by brain damage. There is thus a fertile field in the large number of memory tests which have been developed but not utilized in neuropsychology. In addition, the theoretical concepts that have been developed in the last few years have not been utilized in the construction of clinical memory tests. The primary development here has demonstrated that there are many forms of memory and that these forms of memory are in many cases largely unrelated to each other. Work in this area has been summarized by Russell (1981).

There are in fact many types of memory which need to be taken into consideration in the construction of a new memory test. In summary, some of these types can be described by grouping or comparison of different types of memory. First there is the general process of memory which includes short term memory, a consolidation process into long term memory and finally long term storage. This grouping is equivalent to the neurologists separation of immediate, recent and remote memory. Any battery of memory tests that is developed must measure each of these types of memory.

A second division is that between verbal and figural memory. Again following the concept of lateralization it is quite clear now that in almost all areas there is a distinction between verbal or left hemisphere memory and figural or right hemisphere memory. Since these require measurements of entirely different functions a battery must contain two separate sets of tests to measure these two functions.

Another division that is important is the one more recently proposed between episodic and semantic memory (Russell, 1981). Episodic memory is a memory for particular events which are coded as to time and place. In episodic memory one remembers that a particular thing happened at a particular time and place. Semantic memory only applies to verbal information such as memory of words which have no time or place context. That is we remember the definition of a particular word without remembering where we learned it. This term is primarily verbal. However there is also a nonverbal or figural component such that the person will learn that a particular nonverbal relationship exists. For instance he will know that a particular instrument such as a claw hammer is used for pulling nails but he may not remember where he learned this. Since this non-figural memory exists another term needs to be used. Russell (1981) suggested the term "categorical memory" to cover all types of learning whether verbal or figural.

Another division which is quite traditional is the division between recall and recognition. Recent work by Milner (1978) and others have indicated that there is a real distinction between recall and recognition. The distinction may be quite important since certain mental conditions such as aging appears to effect recall more than recognition.

Finally there is a difference between rote and meaningful memory. Work by Bartlett (1932) in England, a considerable time ago, indicated that when meaningful items are remembered, they are remembered as a general idea and not in specific details. The details are reconstructed in the process of recall. This is quite different from rote memory in which the item is remembered exactly. Some factor analysis have indicated there are different types of memory (Kelley, 1964; Masson, 1977). It is obviously rather important to make a distinction between these types when describing a person's memory or the impairment of memory due to an organic condition.

While some of the types of memory apparently are related to location such as the difference between verbal and figural, others have not been demonstrated to be related to any specific location in the brain. Nevertheless they are important to evaluate when one is considering rehabilitation and maintenance of patients who have organic lesions.

Further research may be able to determine a location for these types of memory.

Thus an important aspect of this discussion is the present clear understanding that there is no such single thing as memory. Instead, there are various memories each of which must be tested separately. To say that a person has a memory deficit, in light of modern theory, is almost a meaningless statement. A person, for instance, may have completely impaired recent memory or long term memory while his immediate memory is intact. Plans for this patient obviously must be quite different depending on the type of memory impairment that the person has. Thus, it is quite evident that "memory" must be tested by a battery of tests.

Development of a New Memory Scale

The time is now right for the development of a new memory scale. There is both the theoretical background and the tests for such a new scale. The requirements of such a battery are fairly well known. First of all many memories need to be tested individually. Thus a battery of tests is required. However it is also obvious that two many memories exist to be tested individually. Consequently, some selection of memories to be tested is required. However, in some cases it may be possible to combine measures so that what appears to be one test may provide several different scores and thus measure different memories. For instance, when material is directly presented the first time and then tested, immediate memory is being tested. If the same material is repeated until the person learns it to a certain criterion, then learning i.e. categorical memory, is being tested and not immediate memory. Thus, the way in which a test is constructed may enable the examiner to measure a number of kinds of memory in one procedure.

From existing theory, a proposed battery of memory tests should cover these types of memory. Short term verbal, short term figural, consolidation of verbal material, consolidation of figural material, storage of verbal episodic memory, storage of figural episodic memory, storage of verbal categorical memory and storage of figural categorical memory. This would cover the major types of memory. But it still does not take into consideration recognition or the memory of meaningful material. The battery could be constructed such that it covered most of these memories individually, but also utilized tests which are found in other neuropsychological testing. For instance, the Digit Span could be used for short term verbal and information for verbal storage presented in the WAIS. Consequently, one would not need to duplicate this test. Ingenuity should be used in combining various tests both within memory and with other neuropsychological tests.

Other psychometric principles, of course, need to be utilized. A battery of tests must be standardized in such a way that it can be compared to other tests and is useful for the examiner. Individual scales should be homogenious. Tests of reliability and validity must be done. Thus, the entire battery can be constructed on a useful and valid basis.

At this time such a test battery has not been developed and it will probably be several years at the earliest before one exists. The question then arises as to what the neuropsychologist can do in such a situation.

Practically at this time, there are really only two possibilities. First, the revised WMS (Russell, 1975) can be utilized and secondly the individual psychologist can construct his own memory scale. One of the best self-constructed memory scales appears to be a memory scale used at the Boston VA Medical Center. However, it is not in published form at this point. A number of people have created their battery by adding certain tests to the Wechsler Memory Scale.

In conclusion, it is quite clear now that the times are propitious for development of new neuropsychological tests based on improved psychometric and neuropsychological theory. This is particularly true for the development of memory tests. There is intense need for good tests, an adequate theory background, and many already developed experimental tests. Thus, the next years should be fruitful for the development of new neuropsychological tests and for those of memory in particular.

References

Bartlett, R.C. Remembering: A study in experimental and social psychology. Cambridge: Cambridge University Press, 1932.

Boll, T.J. The Halstead-Reitan Neuropsychology Battery. In: S.B. Filskov and T.J. Boll (eds) Handbook of Clinical Neuropsychology, New York: Wiley-Interscience, 1981.

Cattell, R.B. The measurement of adult intelligence. Psychological Bulletin, 1943, 40, 153-193.

Erickson, R.C. and Scott, M.L. Clinical memory testing: A review. Psychological Bulletin, 1977, 84(6), 1130-1149.

Golden, C.J. A standardized version of Luria's neuropsychological tests: A quantitative and qualitative approach to neuropsychological evaluation. In: S.B. Filskov and T.J. Boll (eds) Handbook of Clinical Neuropsychology, New York: Wiley-Interscience, 1981.

Goodglass, H. and Kaplan, E. The Assessment of Aphasia and Related Disorders. Philadelphia: Lea & Febiger, 1972.

Guilford, J.P. Psychometirc Methods. New York: McGraw-Hill, 1954.

Guilford, J.P. Nature of Human Intelligence. New York: McGraw-Hill, 1967.

Horn, J.L. Human abilities: A review of research and theory in the early 1770's. Annual Review of Psychology, 1976, 27, 437-485.

Kelley, H.P. Memory Abilities: A Factor Analysis. Psychometirc Monograph No 11, Richmond, Va.: William Byrd P., 1964.

Klein, P. Psychometrics and Psychology. New York: Academic Press, 1979.

Lezak, M.D. Neuropsychological Assessment. New York: Oxford, 1976.

Logue, P. and Wyrich, L. Initial validation of Russell's revised Wechsler Memory Scale: A comparison of normal aging versus dementia. Journal of Consulting and Clinical Psychology, 1979, 47(11), 176-179.

Luria, A.R. The Working Brain. New York: Basic Books, 1973.

Luria, A.R. Higher Cortical Functions in Man (2nd edition). New York: Basic Books, 1980.

Luria, A.R. and Majovski, L.V. Basic approaches used in American and Soviet clinical neuropsychology. American Psychologist, 1977, 32(11), 959-968.

Masson, M.E.J. A multivariate approach to the study of memory and problem solving. MA Thesis: University of Colorado, 1977.

Milner, B. Clue to the Cerbral organization of memory. In: P.A. Buser and Rouger-Buser (eds) Cerebral Correlates of Conscious Experience. Interm Symposium # 6, Elsevier/North Holland: Biomedical Press, 1978.

Nunnally, J.C. Psychometric Theory, New York: McGraw-Hill, 1967.

Prigatano, G.P. The Wechsler Memory Scale is a poor screening test for brain dysfunction. Journal of Clinical Psychology, 1977, 33, 772-777.

Prigatano, G.P. Wechsler Memory Scale: A selective review of the literature. Journal of Clinical Psychology, 1978, 34, 816-832.

Russell, E.W. A multiple scoring method for the assessment of complex memory functions. Journal of Consulting and Clinical Psychology, 1975, 43(6), 800-809.

Russell, E.W. Memory or memories? Paper presented at the 86th Annual American Psychological Association Convention, Toronto, Canada, 1978.

Russell, E.W. The Pathology and Clinical Examination of Memory. In: S.B. Filskov and T.J. Boll (eds) Handbook of Clinical Neuropsychology, New York: Wiley, 1981.

Wechsler, D. A standardized memory scale for clinical use. Journal of Psychology, 1945, 19, 87-95.